

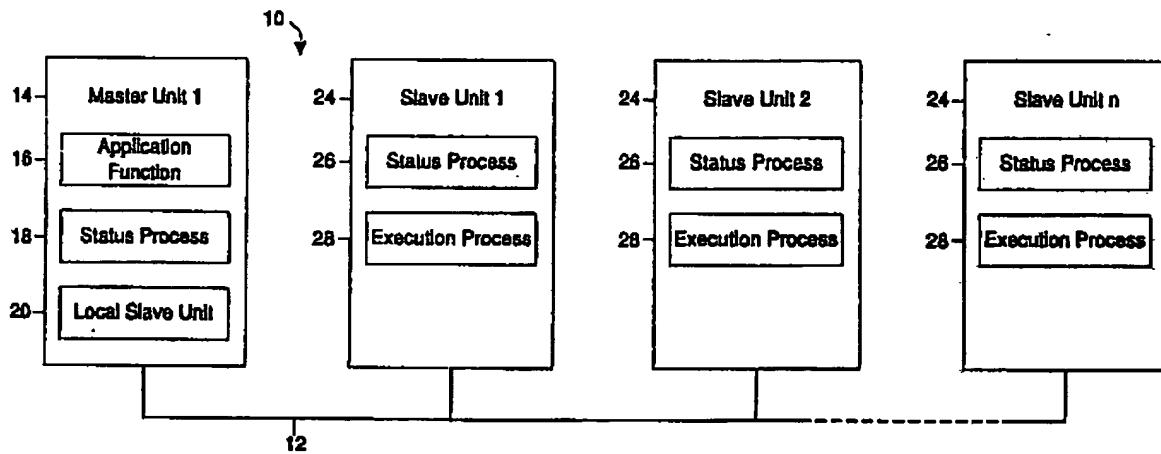


B7

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(31) International Patent Classification 5 : G06F 9/00, 9/302, 9/38 G06F 9/40, 15/56		A1	(11) International Publication Number: <b>WO 93/15457</b> (43) International Publication Date: 5 August 1993 (05.08.93)
(21) International Application Number: PCT/US93/00545 (22) International Filing Date: 21 January 1993 (21.01.93)		(81) Designated States: AU, BR, CA, JP, KR, RU, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Priority data: 829,098 31 January 1992 (31.01.92) US		Published <i>With international search report.</i>	
(71) Applicant: PARALLEL PCs, INC. [US/US]; 1404 Durwood Drive, Reading, PA 19609 (US).			
(72) Inventor: ROSENBERRY, Steven ; 1404 Durwood Drive, Reading, PA 19609 (US).			
(74) Agents: GOLDHAMMER, Joel, S. et al.; Panitch Schwarze Jacobs & Nadel, 1601 Market Street, 36th Floor, Philadelphia, PA 19103 (US).			

## (54) Title: A PARALLEL PROCESSING SYSTEM FOR PERFORMING AN APPLICATION FUNCTION



## (57) Abstract

A parallel processing system (10) performs an application function (16) by a plurality of processing units contained within a single network (12). The parallel processing system includes at least one slave unit (24) connected to a common network and capable of processing a specified amount of data at a particular processing speed. At least one master unit (14), which may include a local slave unit (20), is connected to the common network and initiates an application function to be processed by the slave units. The slave units are polled to determine the processing power of each slave unit, the resources available to each slave unit and the availability of each slave unit. One or more slave units are engaged in processing the application function. If a plurality of slave units are selected, the application function is divided into a plurality of portions of data and each portion is assigned and communicated to a specific slave unit for processing. Predetermined calculations are performed on each portion of data by each slave unit.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SK	Slovak Republic
CI	Côte d'Ivoire	LJ	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republik	MC	Monaco	TG	Togo
DE	Germany	MG	Madagascar	UA	Ukraine
DK	Denmark	ML	Mali	US	United States of America
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

-1-

5

**A PARALLEL PROCESSING SYSTEM FOR  
PERFORMING AN APPLICATION FUNCTION**

**Background of the Invention**

The present invention relates to a parallel processing system and, more particularly, to a parallel processing system for performing an application function 10 by a plurality of processing units contained within a single network.

A typical parallel processing system comprises a plurality of processing units which are interconnected within a single network which together compute the result 15 of a solvable problem. An advantage of the parallel processing system is the increased computing power derived from combining the processing power of several processing units. Typically, many prior art parallel processing units require a known, unchanging parallel processing 20 system to be used. A function which is to be parallel processed is divided and communicated to a plurality of units such that each unit processes a portion of the entire function. Once each processing unit has completed processing its portion of the function, the processed 25 portions of the function are combined to provide a completed function. However, many prior art parallel processing systems arbitrarily divide the application function among the plurality of processing units such that each processing unit completes its portion of the function 30 at a different time. This leads to inefficient use of

-2-

processing time. In addition, many of the processing units in the network typically are not utilized which further lowers the efficiency of the system.

An important aspect to parallel processing systems is the ability of the system to withstand equipment failures in individual processing units and the capability of continued operation. If a parallel processing system is unable to detect a failure in one of its processing units, the system will be unable to complete the processed function. Since the likelihood of a failure occurring during the computation of a given function is significantly high, a parallel processing system which is unable to detect failures is ineffective. On the other hand, if the parallel processing system can detect failures within a single processing unit, but the time necessary to check each processing unit is significant, the fault detect feature will counterbalance any benefit derived by using the parallel processing system.

There is a need for a parallel processing system which is capable of effectively dividing and processing an application function within a minimal amount of time in a processing environment that may change between subsequent executions of the application function. In addition, the parallel processing system should not affect the normal operations of each processing unit. Therefore, when a particular processing unit is not involved in a parallel process, the processing unit may engage in normal processing operations. The parallel processing system should be capable of running on any operating system and should include a plurality of processing units contained within a single network. A

master unit controls the operating flow of the application function and divides the function among a plurality of slave processing units such that each processing unit completes its portion of the function at approximately the 5 same time. Once each slave processing unit completes its portion of the application function, the processed data is transferred back to the master processing unit. In the event that a failure occurs in one of the processing units, the master processing unit reassigns the portion of 10 the function allotted to the failed processing unit to another processing unit such that a completed application function can be performed.

Summary of the Invention

A parallel processing system performs an 15 application function by at least one processing unit contained within a single network. The parallel processing system comprises at least one slave unit. Each slave unit is connected to a common network and is capable of processing a specified amount of data at a particular 20 processing power. At least one master unit is connected to the slave units through the common network. The master unit initiates an application function which is to be processed by the slave units. A polling means polls the slave units to determine the processing power of each 25 slave unit and the resources available to the slave unit. A receiving means receives from each slave unit the processing power and the resources available to each slave unit. A selecting means selects a number of slave units to be engaged in processing said application function. A 30 dividing means divides the application function in a manner defined by the application function into a

plurality of specified portions of data. Each specified portion of data is communicated to a different slave unit for processing. A calculating means performs predetermined calculations as defined by the application 5 function on each specified portion of data by each slave unit. An optional merging means merges each calculated portion of data in a manner defined by the application function at said master unit. The merged data represents the processed application function.

10 A method for performing parallel processing of an application function by at least one processing unit contained within a single network polls a plurality of slave units to determine the processing power of each slave unit. The master unit receives the processing power 15 of each slave unit. The master unit selects a number of slave units to be engaged in processing the application function. The application function is divided in a manner defined by the application function into specified portions of data. Predetermined calculations as defined 20 by the application function are performed on the specified portions of data. The calculated specified portions of data are optionally merged in a manner defined by the application function at the master unit to provide the application function.

25

Brief Description of the Drawings

The foregoing summary, as well as the following detailed description of the preferred embodiment, will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating the 30 invention, there is shown in the drawings an embodiment which is presently preferred, it being understood,

however, that the invention is not limited to the specific methods and instrumentalities disclosed.

In the drawings:

5 Fig. 1 is a general block diagram of a parallel processing system in accordance with the present invention;

10 Fig. 2 is a general block diagram of an application function to be processed by the parallel processing system of Fig. 1;

15 Figs. 3a-3d are flow diagrams depicting the general operation of the parallel processing system of Fig. 1;

Description of Preferred Embodiment

Referring to the drawings, wherein like numerals indicate like elements throughout, there is shown in Fig. 1 a schematic block diagram of a parallel processing system 10 in accordance with the present invention. The parallel processing system 10 contains a plurality of processing units which are contained within a single communications network 12. In the preferred embodiment, the network 12 is a local area network. However, it is to be understood by those skilled in the art that any particular type of local area network can be used or any other suitable means of communications without departing from the scope and spirit of the present invention. Each processing unit is preferably a conventional computer which includes a central processing unit (CPU), memory and file storage and as such will not be described in further detail. The processing unit may 30 be defined as a process or a group of cooperative processes which execute on one or more CPUs. In the

preferred embodiment, each CPU operates under the control of a multi-tasking operating system, such as, but not limited to, UNIX, OS/2 or any other suitable operating system.

5        The parallel processing system 10 includes at least one master unit 14 which stores an application function 16 which is to be executed by the parallel processing system 10. The application function 16 is preferably an application program that performs a specific  
10      function or functions of interest to the unit operator. For example, the application function 16 could relate to word-processing or calculating spread sheets. The master unit 14 also includes a status process 18 which monitors the use of the CPU within the master unit 14 and provides  
15      information relating to the operating system and the resources available to the particular unit. Such information can include the file storage space, the presence of a math co-processor, the amount of the available memory or any other resource information  
20      necessary to determine the suitability of a processing unit to perform a specific application. The master unit 14 includes a local slave unit 20 which is capable of processing a portion of or all of the application function 16. In the preferred embodiment, the local slave  
25      unit 20 uses the same CPU as the master unit 14.

Also contained within the network 12 are one or more remote slave units 24 which are capable of performing at least a portion of the processing of the application function 16. Each slave unit 24 contains a status process 26 which is capable of providing information relating to the resources available to the particular slave unit 24, and the processing power of the slave unit 24. An

execution process 28 executes a portion of the application function once it is distributed to the particular slave unit 24. It is to be understood by those skilled in the art that any unit contained within the network 12 is 5 capable of being designated as the master unit 14 so long as the unit is capable of executing an application function. It is further to be understood that more than one unit contained within the network can be a master unit 14 at the same time. The selection of a slave unit 24 10 used to process a particular application function may depend upon the availability of the slave unit 24 and the power with which the slave unit 24 is capable of processing the application function. In addition, if a particular slave unit 24 is processing its own application 15 function or an application function from a different master unit 24, the slave unit 24 will be unavailable to the polling master unit 14.

Referring to Fig. 2, there is shown a general block diagram of the processing of an application function 16 by the parallel processing system 10. An operator interface 30 communicates data between the application function 16 and a computer operator. A communication may be, but is not limited to, a command with parameters that cause the application function 16 to execute or a 25 graphical operator interface with menus.

A divide function 32 provides a strategy defined by the application function 16 for dividing an application calculate function among multiple units. The divide function 32 typically divides input and output data 30 blocks into smaller ones. The divide function 32 is primarily based on the processing power of each unit as will be described in further detail hereinafter.

-8-

A calculate function 34 generates a desired output through programmed manipulation, as defined by the application function 16, of input data received from the master unit 14. The generated output can be processed 5 data or performance of a specified action such as, but not limited to, data base records, numerical results, graphical displays, printing of reports or control of a process. It is to be understood by those skilled in the art that any application function 16 which is capable of 10 being parallel processed can be generated by the calculate function 34.

A merge function 36 provides a strategy defined by the application function 16 for merging outputs from the different units which execute the calculate function 15 34. The merge function 36 is an optional function and is not necessary for completing the processing of an application function. For certain application functions, such as processing an array of numbers, for example, each processed portion of data may be communicated from the 20 remote slave units 24 to the master unit 14 in a manner that eliminates the need for a merge function..

An abort function 38 selectively discontinues the execution of the calculate function 34 and is part of the fault tolerance processing. Typically, the abort 25 function 38 is determined by the particular calculate function 34 being applied. The abort function 38 is also optional. The presence of the abort function 38 oftentimes provides better performance and more controllable terminations of the calculate function 34.

30 Application processing interface functions (API) 40 provide a means for an application function 16 to access parallel processing services

provided by other slave units. An application function 16 is defined by a task definition. A task definition contains all of the information necessary for the execution of a calculate function 34. Among the 5 information included in the task definition is the location and quantity of any input data required by the calculate function 34, destinations of and quantity of any output data expected to be generated by the calculate function 34, identification and location of the 10 application defined functioning, information concerning minimum resource requirements a unit must have to execute the calculate function 34 and information to estimate the execution time of the calculate function 34 on a given unit.

15 The API functions 40 define the use of and content of the process control data used by the divide function 32 and the merge function 36. The process control data includes the total number of times the divide function 32 or merge function 36 should expect to be 20 executed, the sequence number of the current execution of the divide function 32 or merge function 36, and other information necessary to control the application's divide function 32 or merge function 36. The API functions 40 also define the use of, but not the content of, the 25 application control data. The application control data is used by the divide function 32 and the merge function 34 to manage the progress of their respective executions and the use of any additional system resources required when processing the application function in parallel. The API 30 functions 40 control the allocation of the application control data and its assignment to subtask definitions.

-10-

In determining which slave units 24 will aid in the processing of an application function, the status process 18 or 26 determines the availability and type of resources available to the processing unit on which the 5 status process 18 or 26 is executing. Information necessary for determining the status of a processing unit includes, but is not limited to, the following: the CPU type, the type of operating system, the organization method used by the mass storage device, the unit's base 10 processing power as measured when used to process a single application, the unit's current processing power which is the amount that is currently available for use in processing a task definition, and the amount of free file space that is available for use in processing a task 15 definition. The processing power of a particular remote slave unit 24 or local slave unit 20 is determined by estimating the execution time required by the slave unit to process a given task definition. Typically, the processing power depends upon the slave unit's CPU speed, 20 available memory resources, and the availability of any specialized hardware that may be particularly suited to processing a given task definition.

The status process 18 determines the CPU speed of a particular unit by maintaining a record of its speed 25 over one or more periods of time. The recorded CPU speeds are averaged to provide a CPU speed appropriate for determining a slave unit's processing power for the given task definition.

Slave units 24 compare their available 30 resources as determined by status process 26 with the amount of resources required to complete the application function 16. Only slave units 24 which meet the

requirements of the application function 16 will respond when polled by the master unit 14. Also the slave units 24 must be in an idle state. In the preferring embodiment, the idle state may be determined by: 1) 5 comparing the slave unit's current processing power with its base processing power, and 2) comparing the slave unit's current processing power with the current processing power of the master unit 14 which is polling the slave units 24. However, it is to be understood by 10 those skilled in the art that the idle state may be determined by any other suitable means without departing from the scope and spirit of the present invention. If the slave unit 24 is determined to be busy (not idle), the slave unit 24 will not respond.

15 Referring to Figs. 3a-3d, there is shown a detailed flow diagram depicting the operation of the parallel processing system 10. An application function 16 to be processed by the parallel processing system 10 begins processing on the master unit 14 at block 50. API 20 functions 40 are used to create and define one or more task definitions for the application function 16 in block 52. The task definition is placed on a task list to be processed in a logical manner by an appropriate combination of the master unit 14, the local slave unit 20 25 and one or more of the remote slave units 24.

Once a task list has been formulated, the application function processing is initiated at block 54. In block 56, it is determined whether the application function can be processed in parallel. If the application 30 function 16 cannot be processed in parallel, stand-alone processing may be performed by the master unit 14 at block 58.

-12-

If it is determined that the application function 16 may be processed in parallel, the master unit first determines the amount of resources available to the master unit 14 by checking the status process 18 at block 5 60. The status process 18 includes information, such as the processing power of the local slave unit 20 and the available resources. In block 62, the master unit 14 polls the remote slave units 24 in an attempt to locate available slave units by sending a status request message. 10 The status process 26 of each remote slave unit 24 provides the processing power and resources available for use by the master unit 14. The remote slave units 24 may return a number of possible responses.

If no slave units 24 are available to master unit 14, then the master unit 14 will initiate processing of the application calculate function 34 in block 66. Those slave units 24 which are available to aid in the processing of the application function 16 will send a message back to the master unit 14 in block 68. The slave 20 unit 24 will respond in the affirmative that it has the necessary resources to process the application function 16. Slave units 24 responding in the affirmative are reserved for use by the master unit 14 for a specified period of time or until they receive notification that 25 they have not been selected by the master unit 14. If the slave unit is unavailable, the message may indicate that the particular slave unit is already reserved but not in use by another master unit. Once the master unit 14 has received responses from all responding slave units 24, the 30 master unit 14 must select the slave units which will perform the application function 16 in block 70. The master unit 14 typically selects those slave units 24

which have the most processing power to complete the application function 16. Once the master unit 14 has selected the slave units 24 to process the application function 16, the master unit 14 then notifies all slave units of the selected units 24 in block 76. A slave unit which has not been selected to process the application function 16 and receives the master unit 14 response, will immediately make itself available to other units requiring parallel processing. A slave unit 24 which has been selected by the master unit 14 will commit itself to the master unit 14 provided that it is still available for processing. In the event that a selected slave unit 24 is no longer available for processing, the master unit 14 will either assume that slave unit's processing responsibilities through the fault-tolerance mechanism or select an alternate slave unit for processing. If no remote slaves were selected, the master unit 14 will initiate processing of the application calculate function 34 in block 74.

When the selected slave units 24 receive notification of their selection, bidirectional communications are established between the master unit 14 and each remote slave unit 24 via a separate remote slave thread at block 78. In the preferred embodiment, the remote slave thread is a "lightweight" process executing under the control of the master unit 14. A lightweight process is one that is easily scheduled for execution by the operating system without the overhead of scheduling an actual process. The remote slave thread communicates data, executable software and processing information to and from the master unit 14 and the slave units 24. If at any time during the processing an unrecoverable error

-14-

occurs in the bidirectional communications, the bidirectional communications will be severed. When this occurs, the slave unit 24 is immediately available for use by other master units 14. Subtask definitions assigned to 5 the slave unit 24 with which bidirectional communications were lost will be processed by the master unit 14 according to the fault-tolerance mechanism.

Once the slave units 24 have been notified of the selections made by the master unit 14, subtask 10 definitions for the selected slave units 24 and the local slave unit 20 are generated at block 80 for parallel processing by the slave units. Each subtask definition comprises a portion of the original task definition. The portion of the task definition which is assigned to each 15 subtask definition is dependent upon the execution time required by a specific slave unit to process a portion of the task definition in comparison with the execution times required by the remaining slave units. The percentage of the task definition assigned to each subtask definition is 20 determined at block 82. A divide function 32, provided by the application function 16, is used to define each subtask definition for the application function 16 at block 84. The divide function 32 may deviate from the assigned percentage. In the preferred embodiment, the 25 division of the task definition into subtask definitions is organized such that each subtask definition requires substantially the same amount of processing time. However, it is to be understood by those skilled in the art that the subtasks may contain any arbitrary percentage 30 of the task definition without departing from the scope and spirit of the present invention.

-15-

The percentage of the task definitions assigned to each slave unit 24 is determined in the following manner for all but the last assigned slave unit:

5      
$$\text{Assigned \%} = \frac{(\text{Slave's Current Power}) (100-\text{Total Percentage})}{\text{Total Power of Still Unused Slaves}} \quad (1)$$

For the first slave unit assignment, the total percentage is set to zero. For each subsequent slave unit assignments, the total percentage is equal to the sum of the completed slave unit assignments assigned by the 10 divide function 32. A counter is set to one for the first execution of the divide function 32 and incremented by one for each subsequent divide function 32. The maximum count for the counter is set equal to the number of slave units 24 expected to participate in processing the 15 application function 16. The percentage assigned to the last slave unit is determined by the equation:

$$\text{Assigned to Last Slave Unit \%} = 100-\text{Total Percentage} \quad (2)$$

Once the value of the total percentage is greater than or equal to one hundred, no more calls to the 20 divide function 32 are made. The divide function 32 is expected to return the percentage of the original task definition actually assigned to a subtask definition. The divide function 32 receives the requested percentage of the task definition to be assigned to the subtask 25 definition and the counter values through the process control data as described above. The divide function 32 also has application control data allocated for its use.

Each subtask definition assigned to a remote slave unit 24 is communicated to the assigned slave unit 30 24 via the remote slave thread at block 88. Also, the

-16-

data and executable software necessary for processing the subtask definition are communicated to the slave unit 24 via the remote slave thread at block 90. In addition, in the preferred embodiment, a caching mechanism may be used 5 to store some executable software and some data at the slave unit 24. The caching mechanism eliminates the need for communicating the same information to the remote slave unit 24 during subsequent executions of the application function 16.

10 Once each portion of data has been received by the assigned slave unit 24, each slave unit 24 begins performing the necessary calculations via the execution process 28 at block 92. Prior to performing the calculations, the execution process 28 of each remote 15 slave unit 24 must gain access to all resources required by the calculate function 34 at block 93. Accessing these resources includes locating and gaining permission to access the input, output and constant data assigned to the task definition. If one of the necessary resources cannot 20 be located or accessed, the execution process 28 terminates operation and indicates a system failure.

Once all of the necessary resources are located, the calculate function 34 is executed at block 94. The calculate function uses API functions 40 to 25 locate the data to be processed and to determine the destination of the output data.

While each slave unit 24 is processing the subtask definition or portion of data assigned by the master unit 14 during the calculate function 34, the slave 30 unit 24 monitors the master unit 14 for detection of an abort signal at block 96. A master unit 14 may signal an abort if the local slave unit 20 has already completed the

-17-

assigned subtask definition or in the event that the application function 16 is aborted. If the slave unit 24 receives an abort signal, the slave unit 24 terminates its execution process 28 and waits for further instructions

5 from the master unit 14 at block 100. If the slave unit does not receive instructions from the master unit within a predetermined period of time, the slave unit 24 becomes available to all other master units. If the slave unit 24 does not receive an abort signal, the slave unit 24

10 continues the execution process in block 102.

Typically, a subtask definition is assigned to the local slave unit 20 of the master unit 14 for processing. Once the local slave unit 20 completes the processing of its assigned subtask definition and

15 redefines the subtask definition to represent the actual output data generated by the execution of the application calculate function 34, the master unit 14 checks all slave units 24 to locate those slave units which have not completed their assigned subtask definition at block 128.

20 A slave unit 24 may not complete its subtask definition due to a system error or to a slower than expected processing speed. The master unit 14 selects and assigns one of the incompletely completed subtask definitions for processing by the local slave unit 20 at block 132. If the remote

25 slave unit 24 completes the subtask definition assumed by the local slave unit 20, then the local slave unit 20 aborts the processing of the subtask definition at block 118. The master unit 14 may then assign a different incompletely completed subtask definition from a different slave unit

30 at block 132. If the local slave unit 20 completes the subtask definition prior to the assigned remote slave unit, the master unit 14 signals an abort to the remote

-18-

slave unit to discontinue processing the subtask definition at block 126. The master unit 14 may check the slave units 24 at other times and may also assign subtask definitions to remote slave units.

5 As the execution process 28 for each slave unit 24 completes its assigned subtask definition, the subtask definition is redefined to represent the actual output data that was generated by the execution of the application calculate function 34 and is to be returned to  
10 the master unit 14. The new definition is used to create one or more output response messages at block 108 which are communicated to the remote slave thread at block 110. The remote slave thread uses the output response to redefine the subtask definition at the master unit 14 in  
15 block 112. The actual output data is now communicated from the slave unit 24 to the master 14 via the remote slave thread and the bidirectional communications at block 114.

Once the processing of all the subtask  
20 definitions has been completed including reception of all output data from remote slave units 24, the master unit 14 merges the output data and produces the resulting processed application function 16 at block 134. As each slave unit 24 completes its subtask definition, the slave  
25 unit 24 is available for processing by another master unit 14. At any time during the processing of a subtask definition by a slave unit 24 for a master unit 14, the slave unit 24 is available for local processing. Local processing will have priority over the processing of a  
30 subtask definition for another master unit 14. This may prevent the completion of the subtask definition processing for the master unit 14. In this case, the

-19-

master unit 14 will process that subtask definition through the fault-tolerance mechanism.

The merge function 36 combines the output data from separate subtask definitions into locations specified 5 by the original task definition and releases any system resources allocated during the execution of the divide function 32. The merge function 36 can use API functions 40 to retrieve information needed to determine how the output data is to be merged and what resources should be 10 released. The merge function 36 may take advantage of the process control data and application control data as defined earlier.

The following is a specific example of a type 15 of application function which may be processed by the parallel processing system of the present invention. It is to be understood by those skilled in the art that the following application function is merely exemplary and that any suitable type of application function may be processed by the parallel processing system. One type of 20 application function which may be processed by the master unit 14 is the creation of a document using a word processing software program and including the function of spellchecking the document. In the preferred example, the document to be spellchecked is a lengthy document which 25 would take a considerable amount of time for the master unit to spellcheck alone.

The document is defined in the application function, hereinafter referred to as the spellcheck function, as being capable of being parallel processed. A 30 task definition is created which defines the document to be spellchecked. The master unit 14 polls a plurality of remote slave units 24 contained within a single network to

-20-

determine if there are any available remote slave units 24 to process the document. If no slave units are available, the master unit 14 spellchecks the entire document.

Those remote slave units which are capable of 5 performing the spellcheck function, i.e., which have the available resources and an adequate processing power, will respond to the master unit 14. The maximum number of remote slave units 24 that may be used by the master unit 14 are defined within the spellcheck function. In 10 this particular example, it is assumed that the master unit can reserve up to four remote slave units 24 to complete the spellcheck function. It is to be understood by those skilled in the art that while four slave units is the chosen number of slave units to process the spellcheck 15 function, that fewer slave units or no slave units may be used by the parallel processing system in the processing of the document.

The master unit 14 reviews the affirmative responses from the responding slave units 24 and selects 20 up to four slave units 24 which have the most processing power. Once the slave units 24 have been selected by the master unit 14, the master unit 14 communicates a connect request message to all responding slave units 24 which indicates the chosen slave units and enables the unchosen 25 slave units to be freed for use by other master units 14. In the event that the slave unit 24 is no longer available to the master unit 14, the master unit will reassign that portion of the spellcheck function which was to be assigned to the particular slave unit through the fault 30 tolerance mechanism.

Next, a subtask definition which will contain a portion of the document and the software necessary to

-21-

complete the spellcheck function by a slave units 20, 24 is created for each slave unit 20, 24. The divide function 32 is executed once for each subtask definition to divide the document among the subtask definitions to be 5 processed by the slave units 20, 24. If the processing power of the slave units 24 and the local slave unit 20 are equal, the document can be equally divided among the remote slave units 24 and the local slave unit 20. In this case, since there are five slave units, including the 10 local slave unit 20, the document can be divided into five equal portions. If, for example, the document is 100 pages long, each slave unit will spellcheck 20 pages of the document. If the processing power of the slave units differ, the subtask definitions will contain proportional 15 amounts of the document such that each slave unit can complete its subtask definition at substantially the same time. For example, if the processing power of one of the slave units 24 is twice as much as another slave unit, the more powerful slave unit will receive twice as much of the 20 document to be spellchecked as the less powerful slave unit.

Once each slave unit receives a subtask definition, it also receives the necessary data and executable software for spellchecking the document. The 25 remote slave units 24 and the local slave unit 20 then perform the necessary calculations in order to spellcheck the document. Once the local slave unit 20 has finished spellchecking its portion of the document, the master unit 14 will determine if any of the slave units 24 have not 30 yet completed their assigned subtasks, either due to being executed slower than was expected or a failure occurring at the slave unit. If the master unit 14 detects an

-22-

incomplete subtask definition, the master unit 14 will assign the subtask definition to the local slave unit 20 to spellcheck that portion of the document. If the slave unit 24 finishes spellchecking the portion of the document 5 prior to the local slave unit 20, the master unit 14 will abort spellchecking that portion of the document being performed by the local slave unit 20. If the local slave unit 20 completes spellchecking the document prior to the assigned remote slave unit 24, the master unit will abort 10 the spellchecking of the document being performed by the remote slave unit 24.

As each remote slave unit 24 completes the subtask definition, the processed subtask definition is redefined to represent the actual output data generated by 15 the calculate function 34 which in this example may be a list of misspelled words and their locations within the document. One or more output responses are communicated by the remote slave to the master unit 14 via the remote slave thread. As each output response is received by the 20 master unit 14, the subtask definition at the master unit 14 is also redefined and the actual output data is received from the remote slave unit 24. The master unit 14 combines all of the output data received from the remote slave units 24 and the output data generated by the 25 local slave unit 20 to obtain a completed list of errors within the spellchecked document. The list of errors within the spellchecked document can then be further processed by the master unit 14 as may be further defined in the application function.

30 From the foregoing description, it can be seen that the present invention comprises a flexible parallel processing system for performing an application function

-23-

by at least one processing unit contained within a single network. It will be appreciated by those skilled in the art that changes could be made to the embodiment described above without departing from the broad inventive concept 5 thereof. It is understood, therefore, that this invention is not limited to the particular embodiment disclosed, but it is intended to cover all modifications which are within the scope and spirit of the invention as defined by the appended claims.

-24-

CLAIMS

1. A parallel processing system for performing an application function by a plurality of processing units contained within a single network, the 5 parallel processing system comprising:
  - at least one slave unit connected to a common network and being capable of processing a given amount of data specified by the application function at a particular processing power;
- 10 at least one master unit connected to said slave units and said common network, said master unit initiating an application function which is to be processed by said slave units;
- 15 polling means for polling said slave units and master unit to determine the processing power of each slave unit and the master unit, the resources available to said slave units and said master unit, and the availability of said slave units;
- 20 receiving means for receiving from each slave unit and master unit the processing power and resources available to each slave unit and master unit;
- 25 selecting means for selecting one or more slave units to be engaged in processing said application function;
- 30 dividing means for dividing said application function in a manner defined by the application function into a plurality of specified portions of data, each specified portion of data being assigned and communicated to a specific slave unit for processing; and
- calculating means for performing predetermined calculations defined by the application

function on each said specified portion of data by each slave unit.

2. A parallel processing system according to claim 1, further including merging means defined by the 5 application function for optionally merging each calculated portion of data at said master unit, said merged data representing said processed application function.

3. A parallel processing system according to 10 claim 1 wherein said polling means further includes estimating means for estimating the processing power of each slave unit.

4. A parallel processing system according to 15 claim 1, wherein said common network is a local area network.

5. A parallel processing system according to claim 1, wherein said master unit can be any one of the processing units located within the common network.

6. A parallel processing system according to 20 claim 5, wherein said master unit includes a local slave unit.

7. A parallel processing system according to 25 claim 6, wherein said local slave unit performs at least a portion of the processing of the application function.

8. A parallel processing system according to claim 1, further comprising fault tolerance means for determining the present status of each slave unit and for detecting an uncompleted calculated portion of data.

9. A parallel processing system according to 30 claim 1, further comprising fault tolerance means for determining the present status of each slave unit and for detecting a fault occurring at the slave unit.

-26-

10. A parallel processing system according to claim 8, further comprising data assigning means for assigning the portion of data which has been originally assigned to said remote slave unit which is still 5 calculating the portion of data to another slave unit for performing the calculation functions on the assigned data.

11. A parallel processing system according to claim 9, further comprising data assigning means for assigning the portion of data originally assigned to a 10 failed remote slave unit to another slave unit to perform the predetermined calculations on the assigned data.

12. A parallel processing system according to claim 1 wherein said master unit is capable of processing a particular application function as a single execution 15 process.

13. A method for performing parallel processing of an application function by a plurality of processing units contained within a single network, said processing units including at least one slave unit, each 20 slave unit being capable of processing a specified amount of data defined by the application function at a particular processing power, and at least one master unit contained within the network, said master unit initiating the application function which is to be processed and 25 including a local slave unit, said method comprising the steps of:

- a) estimating the processing power of each slave unit;
- b) polling said slave units to determine 30 the processing power and availability of each slave unit;
- c) receiving the processing power and availability of each slave unit;

-27-

d) selecting a number of slave units to be engaged in processing said application function;

5 e) dividing said application function in a manner defined by the application function into specified portions of data;

f) assigning each specified portion of data to a specific slave unit; and

10 g) performing predetermined calculations defined by said application function on each said specified portion of data.

14. A method according to claim 13 further comprising the step of:

15 h) merging, as defined by the application function, the calculated specified portions of data at said master unit to provide the processed application function.

15. A method according to claim 13, further comprising the steps of:

20 i) determining which remote slave units have not completed calculating their assigned portion of data; and

j) assigning the specified portion of data originally assigned to the remote slave unit to another slave unit.

25 16. A method according to claim 13, further comprising the steps of:

k) determining which remote slave units have failed; and

30 l) assigning the specified portion of data originally assigned to the failed remote slave unit to another slave unit.

-28-

17. A parallel processing system for performing an application function by a plurality of processing units contained within a single network, the parallel processing means comprising:

- 5 at least one remote slave unit, each remote slave unit being interconnected to other remote slave units by a common network and being capable of processing a specified amount of data defined by the application function at a particular processing power;
- 10 a master unit connected to said remote slave units and said common network, said master unit initiating an application function which is to be processed by said remote slave units;
- 15 retrieving means for retrieving from said slave units and said master unit the processing power of each remote slave unit and master unit, the resources available to each remote slave unit and master unit, and the availability of each remote slave unit;
- 20 selecting means for selecting a number of remote slave units to be engaged in processing said application function; and
- 25 dividing means for dividing said application function in a manner defined by said application function into specified portions of data, each specified portion of data being assigned to a specific remote slave unit for processing.

18. A parallel processing system according to claim 17, wherein said master unit includes a local slave unit.

30 19. A parallel processing system according to claim 18, wherein said local slave unit receives a specified portion of data for processing.

20. A parallel processing system according to  
claim 17, further comprising calculating means for  
performing predetermined calculations defined by said  
application function on said specified portions of data by  
5 said remote slave units and said local slave unit.

21. A parallel processing system according to  
claim 20, further comprising optional merging means  
defined by the application function for merging each  
calculated portion of data at said master unit, said  
10 merged data representing the processed application  
function.

22. A parallel processing system according to  
claim 17, wherein said master unit can be any processing  
unit located within the common network.

15 23. A parallel processing system according to  
claim 21, further comprising fault tolerating means for  
detecting at the completion of a given event those remote  
slave units which are still calculating a portion of data.

24. A parallel processing system according to  
20 claim 23, further comprising data assigning means for  
assigning the portion of data which has originally been  
assigned to said remote slave unit which is still  
calculating the portion of data to another slave unit for  
performing the calculating functions on the assigned data.

25 25. A parallel processing system according to  
claim 21, further comprising fault tolerating means for  
detecting at the completion of a given event those remote  
slave units which have failed.

26. A parallel processing system according to  
30 claim 25, further comprising data assigning means for  
assigning the portion of data originally assigned to a

-30-

**failed remote slave unit to another slave unit to perform  
the predetermined calculations on the assigned data.**

1 / 6

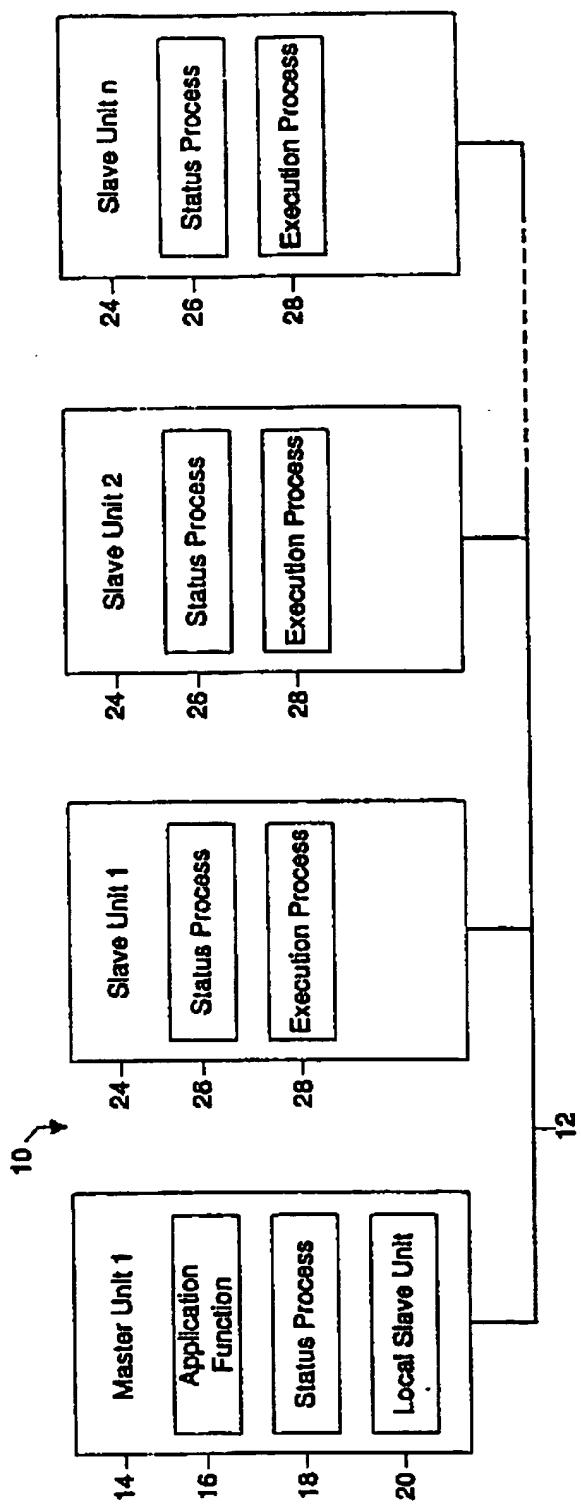


Figure 1

2 / 6

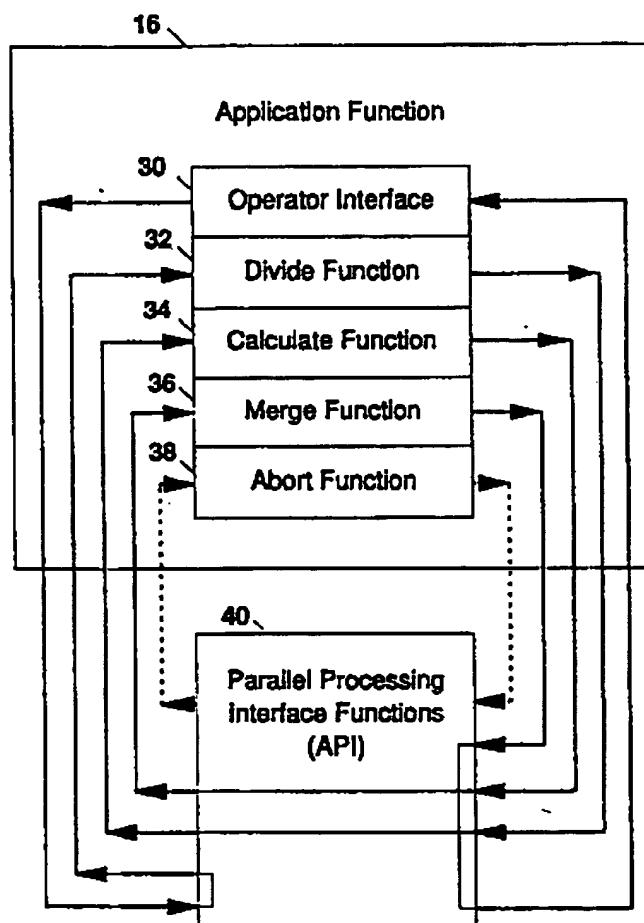


Figure 2

3 / 6

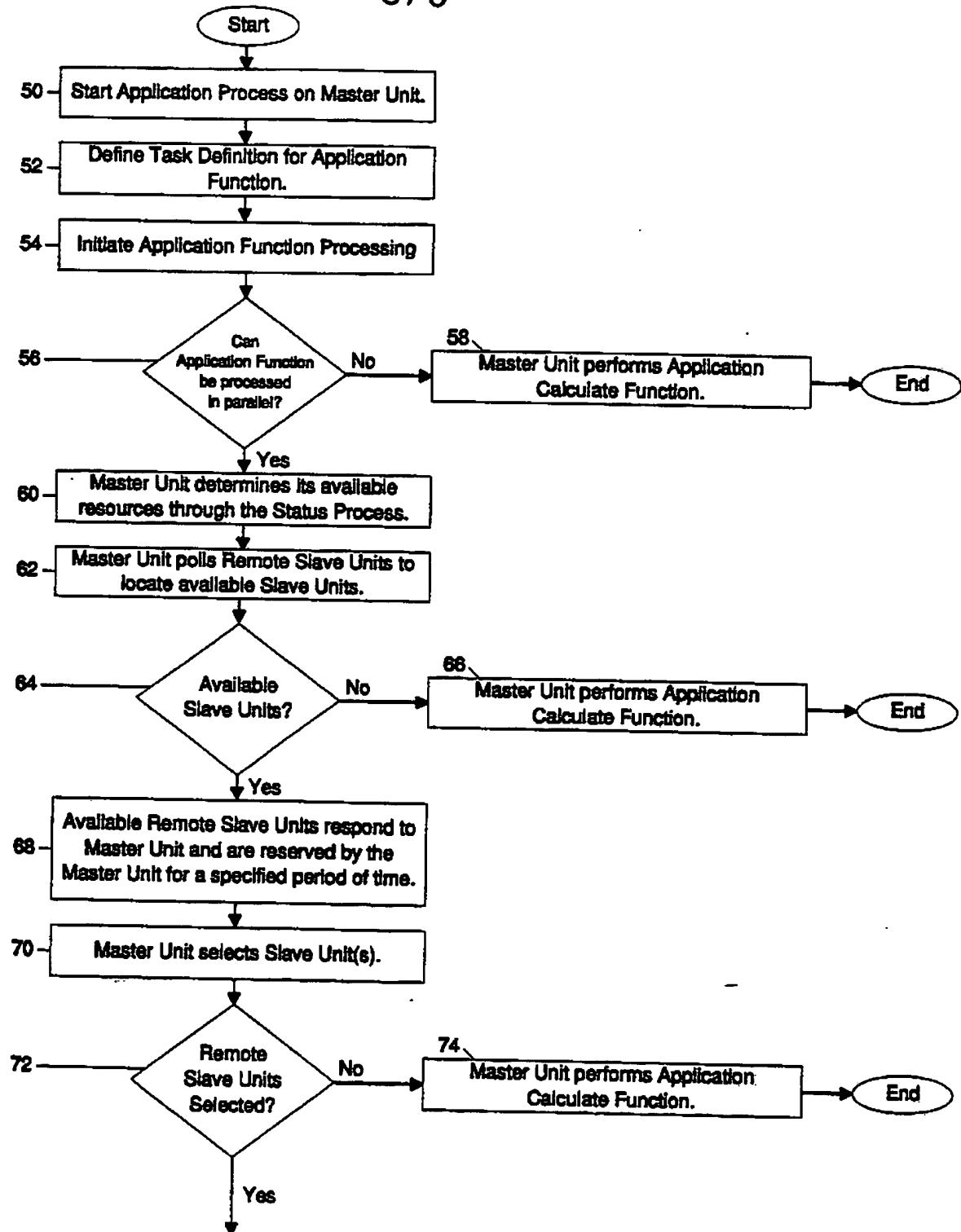


Figure 3a

SUBSTITUTE SHEET

4 / 6

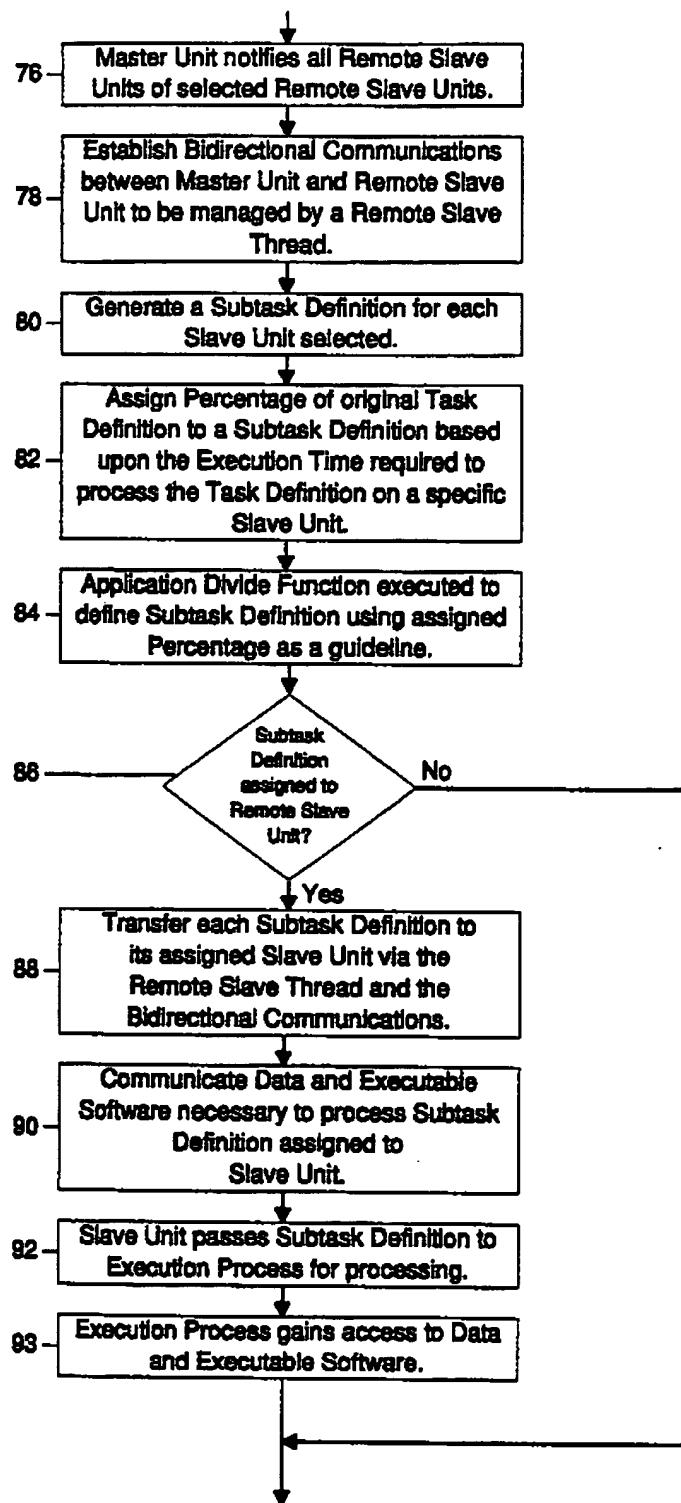


Figure 3b

SUBSTITUTE SHEET

5 / 6

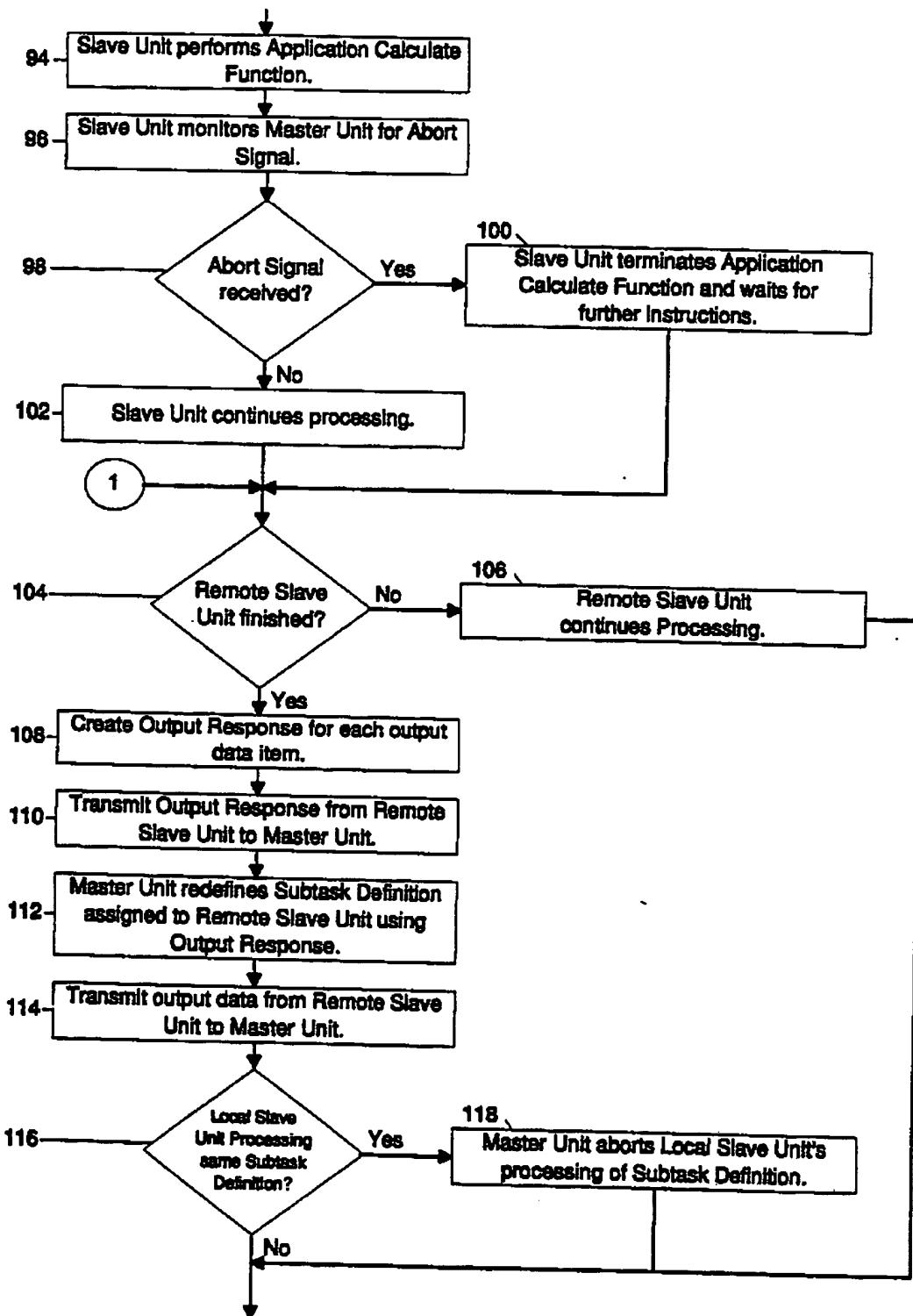


Figure 3c

SUBSTITUTE SHEET

6 / 6

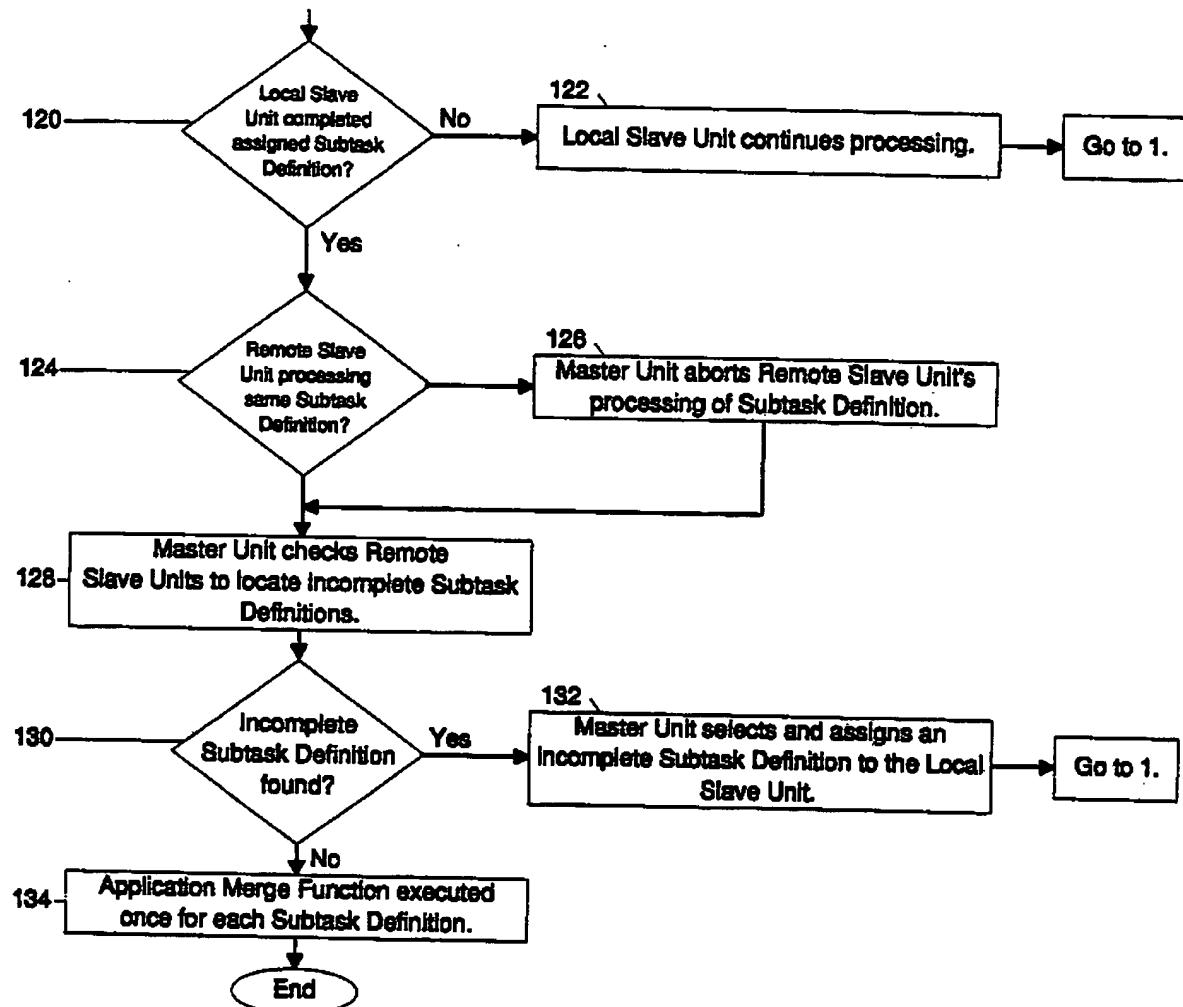


Figure 3d

SUBSTITUTE SHEET

## INTERNATIONAL SEARCH REPORT

PCT/US93/00545

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :G06F 9/00, 9/302, 9/38, 9/40, 15/56  
 US CL :395/800,650,575,375,200

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 395/800,650,575,375,200

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

U.S. Patent Office, Automated Patent System (APS), Polling, LAN, Simultaneous or Concurrent processing, master and slave processing, Bus, Tasks, Jobs, programs, Splitting or dividing, assigning

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US,A, 5,003,464 (Ely et al.) 26 March 1991 See column 2, line 65- column 4, line 13.	1-26
Y	US,A, 5,031,089 (Liu et al.) 09 July 1991 See column 7, lines 13-41.	1-26
Y	US,A, 5,056,000 (Chang) 08 October 1991 See column 2, lines 3-32.	1-26
Y	US,A, 4,769,771 (Lippmann et al.) 06 September 1991 See column 2, line 8- column 5, line 17.	1-26
A	US,A, 4,920,487 (Baffes) 24 April 1990 See column 4, line 10- column 6, line 21.	1-26

Further documents are listed in the continuation of Box C.  See patent family annex.

* Special categories of cited documents:	
*A* document defining the general state of the art which is not considered to be part of particular relevance	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*E* earlier document published on or after the international filing date	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*O* document referring to an oral disclosure, use, exhibition or other means	*Z* document member of the same patent family
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search	Date of mailing of the international search report
12 MARCH 1993	28 APR 1993
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer /s/ ALPESH M. SHAH <i>Alpesh M. Shah</i> Telephone No. (703) 305-9698
Faxsimile No. NOT APPLICABLE	

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US93/00545

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,E	US,A, 5,193,179 (Laprade et al.) 09 March 1993 See column 1, line 65- column 2, line 64.	1-26
Y	US,A, 4,980,824 (Tulpule et al.) 25 December 1990 See column 2, line 3- column 6, line 44.	1-26